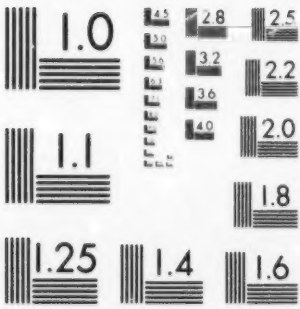

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DOCUMENT RESUME

ED 180 194 FL 010 676

AUTHOR Backman, Jarl
TITLE Sentence Production with Homonymous and Polysemous Words. Educational Reports Umea, No. 18.
INSTITUTION Umea Univ. (Sweden).
SPONS AGENCY Swedish Council for Social Science Research, Stockholm.
PUB DATE 78
NOTE 19p.
AVAILABLE FROM Department of Education, University of Umea, S-901 87 Umea, Sweden (free)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Students; *Language Processing; Language Research; *Linguistic Competence; *Nominals; Psycholinguistics; *Semantics; Swedish; *Verbs
IDENTIFIERS *Homographs; *Polysemy; Swedes

ABSTRACT

Forty Swedish university students produced sentences from homographs that could be interpreted either as verbs or nouns. The words also varied in degree of polysemy (multiple meaning). The results indicated that the subjects prefered verb productions when the words were grouped according to objective frequency. This was more evident when the choice of syntactic category was related to the word's polysemous character. The results were explained by a presumed easier access to and a larger number of entries in the subjective lexicon for verbs, caused by their relational character and context dependence. The nouns, on the other hand, are semantically more stable and less accessible because of a higher degree of polysemy. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

educational reports umeå

no 18 1978

ED180194

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Umeå Univ.
Dept. of Education

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL010 676

Backman, J. Sentence production whith homonymous and polysemous words. Educational Reports, Umeå, no 18, 1968. ×

ABSTRACT

Forty subjects produced sentences from homonymous words which could be interpreted either as verbs or nouns. The words also varied in degree of polysemy. The results indicated that the subjects prefered verbproductions when the words were grouped according to objective frequency. This was more evident when the choice of syntactic category was related to the words polysemous character. The results were explained by a presumed easier access and a larger number of entries in the subjective lexicon for verbs, caused by their relational character and context dependence. The nouns on the other hand are semantically more stable and an increase in degree of polysemy restricts access to a specific interpretation.

× This research has been supported by The Swedish Council for Social Research.

Many psycholinguists today agree on the theoretical standpoint that comprehension of words and sentences have an internal representation beyond linguistic input and that it is of a more abstract nature. The most extensive and operational formulation of such a theoretical position is presented by Schank in "Conceptual Dependency Theory" (Schank, 1972). There is, according to Schank, a conceptual basis which is interlinguistic and which consists of a set of primitive concepts and conceptual relations. The process of comprehension is said to be a transformation of linguistic expressions in this (non-verbal) conceptual basis providing words and sentences whith meaning.

Another theoretical starting-point, allowing more room for lexical information, is the theory that words get their meaning via relations to other words. These so-called associative network thecries of a more extensive kind have been put forward by for example Anderson & Bower (1973) and Norman & Rumelhart (1975). The nodes in the theory of Noman & Rumelhart (1975) consist of concepts (often words) forming complex structures of denominate relations. These relations provide the words with meaning.

No matter which theoretical standpoint is taken regarding word comprehension, it seems that the syntactic categories verbs and nouns play the most important roles. It is for example worth noting that

the conceptual categories in the theory of Schanks (1972) are based on verbs. The verbs have approximately the same status in the network theories. The relations which carry meaning are most often found in the verbs. These two theories are both based on Fillmore's (1968) observations that the classification of verbs provides a sentence with its principal meaning. Chafe (1970) and Miller (1972) also adhere to such a view-point:

"Verbs, on the other hand, seem to be the real motor behind the critically important process on predication..." (Miller, 1972, p 338)

The verb is a relational verbal item, which often determines the representation of other items in a clause. But for simple word comprehension the noun is, however, the syntactic category which has been most exhaustively tested. There are no comparative studies between the categories and according to Johnson-Laird (1974) it is (on the level of the sentence),

"...difficult to establish whether the nouns are more salient than the verbs, or the verb is more salient than the nouns". (Johnson-Laird, 1974, p 145).

Gibson & Levin (1975) pointed out that there are no studies in this matter, but they still venture to formulate the hypothesis that verbs should have lower recognition thresholds since they make up the heart of a sentence.

One way of illustrating a possible difference between nouns (N) and verbs (V) is by letting test subjects produce sentences from homonymous words. Homonyms are words with two or more distinct meanings, for example the word NUMBER which can be both a verb and a noun. A homonym with two meanings is then treated as activating different semantic representations. The choice between N and V can then, in the production of sentences, be taken as an indication of ease or difficulty of comprehension.

A number of variants or groups of meaning can be discerned within one meaning of a word form. Such a word is said to display polysemy or carry multiple meanings. A separate semantic representation is also adopted for each variant of meaning. The hypothesis that words with multiple meanings have lower "lexical decision latency" than non-homographs (Rubenstein, Garfield & Millikan, 1970) has been empirically verified since the former have several entries in a semantic memory. This latency has also been proved to be a function of the number of groups of meaning (Jastrzembski & Stanners, 1975). These experiments implicate the hypothesis that it is not the syntactic categories verbs or nouns <u>per se</u> which determine a choice between them but the number of groups of meaning. It then follows that if subjects are asked to produce a sentence from a homonym they will chose the category which has the largest number or groups of meanings. If, however, you

disregard the polysemy and only observe the objective frequency of the words it has earlier been noted (Backman, 1978 a) that subjects preferably produce verbs on lower frequency levels and nouns on higher levels. This relation was found valid for estimations of comprehension. The aim of the experiment below is to find out which syntactic category a person chooses in producing sentences when the initial word can be interpreted both as a verb and a noun.

METHOD

Subjects

The subjects were 40 students at the Umeå School of Education. The experiment was carried out during regular lesson time and the subjects were not paid for their participation.

Material and procedure

Thirty different homonyms, which could be both nouns and verbs were chosen from Allén (1970; 1971) through stratified sampling. The total frequency was not allowed to exceed 150, since higher frequencies have a tendency to give priority to the use of only one of the categories nouns-verbs. Information about the number of groups of meaning for each syntactic category was taken from "Ord för ord - Svenska synonymer och uttryck" (Word for word - Swedish Synonyms and Expressions) (1964). Table I presents the different words as well as the quota nouns/verbs for both frequency and variant of meaning.

All the words were given on two sheets of paper and for each word the test subjects were instructed to produce one sentence with the given word.

RESULTS

A simple calculation of the proportion N-choices and V-choices totally over frequencies results in an almost identical distribution of the choices: 0.43 and 0.57, respectively. If, however, the total frequency is grouped in a low, middle, and high category the results will be as in Table II. The division into frequency levels is made so that the total number of words (30) has been split up into groups of 10 in each total frequency. The dependent variable is here presented as a quota between N and V, and a quota < 1 means that the sentence productions are dominated by verbs and > 1 that they are dominated by nouns. The quotient < 1 means, on all three frequency levels, that the subjects have chosen verb productions to a larger extent than noun production. Produced relations between the two categories do not covary with total frequency, which is shown by the product moment coefficient: $\underline{r} = .31$, $\underline{t}\ (28) = 1.95$, $\underline{p} > .05$.

Table III shows the result of an analysis based on the quota between number of multiple meanings (Mm) for every word category. Quotas < 0.05 have been classified as verb dominant and > 1.5 as noun dominant. A quota between these values has been classified as neither verb nor noun dominated. The real quotas for these three levels have also been determined and they are

Table I. The words for sentence production, their total frequency (F), frequency per syntactic category (F N/V) and number of meanings per category (Mm N/V).

Word	F	F N/V	Mm N/V
RÄNNA	2	1/1	2/2
GALA	2	1/1	1/3
LÄCKA	4	1/3	2/1
GLÄNTA	5	3/2	1/2
VÄNTA	5	3/2	1/1
HACKA	7	2/5	2/2
SKENA	8	5/3	1/2
SPIRA	9	6/3	5/2
HÄLLA	10	1/9	1/2
FLAGGA	13	11/2	1/2
FLAMMA	17	3/14	2/2
HETTA	20	19/1	2/2
STADGA	25	15/10	2/3
GÖMMA	37	4/33	2/3
STICKA	45	5/40	3/6
SKO	49	48/1	2/2
SPRINGA	51	2/49	1/3
HYRA	55	18/37	4/2
SKALA	58	51/7	4/2
KLIPPA	62	15/47	2/4

Table I

cont.

Word	F	F N/V	Mm N/V
STRÄCKA	70	21/49	1/6
SKÄRPA	71	27/44	6/4
SÄNKA	73	17/56	2/2
TÄMMA	78	37/41	3/2
HÄLSA	85	32/53	1/3
VILA	85	14/71	1/5
STYRKA	108	97/11	4/2
BLOMMA	133	96/37	3/1
SKADA	140	78/62	3/4

Table II. The homonyms grouped in total frequency with mean frequency (M_f) transformed into a subjective frequency scale (SFI) and produced quotas nouns/verbs (N/V).

Frequency level	Low	Middle	High
f-range	2-10	10-58	62-140
M_f	6	37	91
SFI	48	56	60
N/V	0.44	0.56	.30
n	10	10	10

presented under the heading $\underline{N}/\underline{V}_{obj}$. There is almost an identical distribution in the choice of syntactic category for the lowest $\underline{N}/\underline{V}$-level $\underline{N}/\underline{V}_{emp}$ = 0.90 where the verbs accordingly dominate objectively ($\underline{N}/\underline{V}_{obj}$ = 0.39). ($\underline{N}/\underline{V}_{obj}$ = 0.39) the result is a weak verb preference which on the highest $\underline{N}/\underline{V}$-level becomes a clearly marked verb preference.

Table III. Three quota level divisions with empirically obtained ($\underline{N}/\underline{V}_{emp}$) and objective ($\underline{N}/\underline{V}_{obj}$) relations between N and V and the difference between them. The number of words is indicated by $\underline{n}$.

$\underline{N}/\underline{V}$	< 0.5	0.5 < $\underline{N}/\underline{V}$ < 1.5	> 1.5
$\underline{N}/\underline{V}_{emp}$	0.90	0.68	0.68
$\underline{N}/\underline{V}_{obj}$	0.39	0.93	2.07
Diff	+0.51	-0.25	-1.39
$\underline{n}$	12	11	7

The mean frequency for the three quota levels was determined to check if the result is a function of frequency and these mean frequencies are presented in Table IV. This table also indicates the transformation of these values into SFI (Carrol, 1970), which is a subjective frequency scale providing a good psy-

Table IV. The mean total frequency of the three quota levels $\underline{M}_f$ and corresponding transformations into SFI.

N/V	< 0.5	0.5 < N/V < 1.5	> 1.5
$\underline{M}_f$	36	35	73
SFI	56	55	59

chological description of the perception of the frequency of occurrence for words (see e.g. Backman, 1976). The difference between the extreme quotas (N/V < 0.5 and N/V > 1.5) in Table III correspond only to three units in SFI. The grouping of meaning does not seem to follow the objective frequency count and the suspicion of contamination between these variables can accordingly be dismissed.

In an experiment testing if individuals can discriminate between numbers of multiple meanings (Backman, 1978 b) a limit was found to go between 2 and 3 variants. In consequence, a differentiation was made of the different quotas. One group consisted of quotas of which the number of variants in both nouns and verbs was < 3, which it was not supposed to be possible to differentiate. The others were brought to a separate group, where one or both of the syntactic categories thus exceeded the number 2. If we for

example take the word "hyra" the quota is $\underline{N}/\underline{V}$ = 4/2. The noun form has 4 $\underline{Mm}$ and the verb form has 2. It is supposed that the former can be discriminated in the noun position since the number 4 exceeds the discrimination limit 3. It is here supposed that the category results in at least 2 representations. The verb, on the other hand, is only supposed to activate one representation of meaning. For the former group we obtained $\underline{N}/\underline{V}_{emp}$ = 0.67 with $\underline{N}/\underline{V}_{obj}$ = 0.74 an almost negligible difference. This group is, however, interesting as it is assumed that only the homonymy, without interference of polysemy, is active. The subjects are here only expected to be able to perceive the two main categories. If the most elevant words in the group are examined, i.e. those words where the quota is exactly 1, the value 0.59 is obtained for $\underline{N}/\underline{V}_{emp}$, i.e. a clear verb preference. The calculation is made on 8 words where the relations of polysemy in all cases are 2/2 except one which is 1/1.

In the "discrimination group" $\underline{N}/\underline{V}_{emp}$ was found to be 0.85 and the corresponding $\underline{N}/\underline{V}_{obj}$ = 1.19, i.e. a verb preference. If a subgrouping is made in the latter group, i.e. $\underline{N}/\underline{V} < 0.5$ and $\underline{N}/\underline{V} > 1.5$, the quota 0.73 is obtained for verb dominating variants ($\underline{N}/\underline{V}_{obj}$ = 0.33) and the quota 0.68 for the noun dominated ones ($\underline{N}/\underline{V}_{obj}$ = 2.07). This is illustrated in Table V. The tendency to chose verbs instead of nouns is now unmistakable

Table V. Objective and epirical quotas (N/V) and their differences for various numbers of multiple meanings (Mm). The division of the latter into different quota relations (N/V) is illustrated at the bottom on the table.

	N/V_{obj}	N/V_{emp}	$Diff_{emp-obj}$	n
Mm < 3	0.74	0.67	- 0.07	14
Mm > 3	1.19	0.85	- 0.34	16
N/V < 0.5	0.33	0.73	0.40	6
N/V > 1.5	2.07	0.68	1.39	7

and this is most clearly seen when the objective relation between numbers of multiple meanings is noun dominated. When the latter category objectively is twice as big as the number of verb variants, the subjects will never the less produce almost twice as many verb clauses.

DISCUSSION

To sum up, the experiment has demonstrated clear verb preferences when the sentence productions are related to the variable objective frequency. They are, however, not systematically related to frequency, but seem rather to be constant and independent of frequency level within the examined range. This is also emphasized

by the fact that there was no co-variation between frequency and syntactic category production when it was determined via the product moment coefficient.

If, however, the polysemy of the syntactic categories is taken into account, the conslusion will be more systematic. When a discrimination limit for the perception of numbers of multiple meanings is taken into consideration this will lead to clear verb preferences, especially when the numbers of meanings of the nouns exceed the limit. This can perhaps be explained by the status of the verbs as relation terms and their relative context dependence. The nouns are more stable semantically and an increase in the number of variants of meaning restricts the immediate access to a specific interpretation. This access may become more complicated since different variants of meaning will then be competing with each other. The verbs are more unspecified and a change in polysemy will have no effect as long as a context is not introduced. The verbs have, in more distinct theoretical terms, a larger and more indefinite number of entries in the subjective lexicon of an individual.

REFERENCES

Allen, S. Frequency Dictionary of Present-Day Swedish Based on Newspaper Material. 1. Graphic Words Homograph Components. Stockholm: Almqvist & Wiksell, 1970.

Allen, S. Frequency Dictionary of Present-Day Swedish Based on Newspaper Material. 2. Lemmas. Stockholm: Almqvist & Wiksell, 1971.

Backman, J. Some common word attributes and their relation to objective frequency counts. Scandinavian Journal of Educational Research, 1976, 20, 175-187.

Backman, J. Word comprehension as function of syntactic category of content words. Educational Reports Umeå, 1978, no 15. (a).

Backman, J. Perception of polysemi. Educational Reports Umeå, 1978, no 16, (b).

Carroll, J.B. An alternative to Juilland's usage coefficient for lexical frequencies, and a proposal for a Standard Frequency Index (SFI). Computer Studies in the Humanities and Verbal Behavior, 1970, 3, 61-65.

Chafe, W.L. Meaning and the Structure of Language. Chicago: University of Chicago Press, 1970.

Fillmore, C.J. The case for the case. In E. Bach and R.T. Harms (Eds), Universals in Linguistic Theory. New York: Holt, Rinehart & Winston, 1968.

Gibson, E.J., & Levin, H. The psychology of Reading. Cambridge, Mass.: The MIT Press, 1975.

Jastrzembski, J.E., & Stanners, R.F. Multiple word meanings and lexical search speed. Journal of Verbal Learning and Verbal Behavior. 1975, 14, 534-537.

Miller, G.A. English verbs of motion: a case study in semantics and lexical memory. In A.W. Melton and E. Martin, Coding Processes in Human Memory. New York: V.H. Winston & Sons, 1972.

Miller, G.A., & Johnson-Laird, P.N. Language and Perception. Cambridge: Cambridge University Press, 1976.

Norman, D.A., & Rumelhart, D.E. Explorations in Cognition. San Francisco: Freeman, 1975.

Rubenstein, H., Garfield, L., & Millikan, J.A. Homographic entries in the internal lexicon. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 487-494.

Schank, R. Conceptual Information Processing. Amsterdam: North Holland, 1975.

EDUCATIONAL REPORTS, UMEÅ UNIVERSITY AND UMEÅ SCHOOL OF EDUCATION

Earlier reports in this series:

1973

1. READABILITY OF STAVES WITH DIFFERENT LINE AND NOTE DISTANCES.
Jarl Backman and Allan Lindahl

2. LITERACY AND SOCIETY IN A HISTORICAL PERSPECTIVE - A CONFERENCE REPORT.
Egil Johansson (Ed.)

3. THEORETICAL PROBLEMS IN CONSTRUCTION OF CRITERION-REFERENCED TESTS.
Ingemar Wedman

4. RELIABILITY, VALIDITY, AND DISCRIMINATION MEASURES FOR CRITERION-REFERENCED TESTS.
Ingemar Wedman

1974

5. DROPOUT AND PASS-RATE IN HIGHER EDUCATION - ARE THEY USEFUL MEASURES OF EFFICIENCY?
Inga Elgqvist-Saltzman

6. SOME COMMON WORD ATTRIBUTES AND THEIR RELATIONS TO OBJECTIVE FREQUENCY COUNTS.
Jarl Backman

7. ON TEST-WISENESS AND SOME RELATED CONSTRUCTS.
Ingvar Nilsson and Ingemar Wedman

1975

8. THE OCCURRENCE OF TEST-WISENESS AND THE POSSIBILITY OF INDUCING IT VIA INSTRUCTION - A cross-sectional study.
Ingvar Nilsson

1976

9. THE FUNCTION OF GROUP-SIZE AND ABILITY LEVEL ON SOLVING A MULTIDIMENSIONAL COMPLEMENTARY TASK.
Thor Egerbladh

10. READING COMPREHENSION AND PERCEIVED COMPREHENSIBILITY OF LEXICAL DENSITY AT DISCOURSE AND SENTENCE LEVEL.
Jarl Backman

11. EFFECTS OF GROUP LEARNING IN GRADE FOUR.
Thor Egerbladh

1977

12. THE HISTORY OF LITERACY IN SWEDEN. In comparison with some other countries.
Egil Johansson

13. LITERACY SCHOOLS IN A RURAL SOCIETY. A Study of Yemissrach Dimts Literacy Campaign in Ethiopia.
Margareta Sjöström and Rolf Sjöström

1977

14. TASK DIFFICULTY ABILITY LEVEL AND GROUP SIZE: A TEST OF THE THEORY OF SOCIAL DECISION SCHEMES.
Thor Egerbladh.

1978

15. WORD COMPREHENSION AS FUNCTION OF SYNTACTIC CATEGORY OF CONTENT WORDS
Jarl Backman

16. PERCEPTION OF POLYSEMY
Jarl Backman

17. SEMANTIC VARIABILITY AND WORD COMPREHENSION
Jarl Backman